Most Permissive Boolean Networks in practice

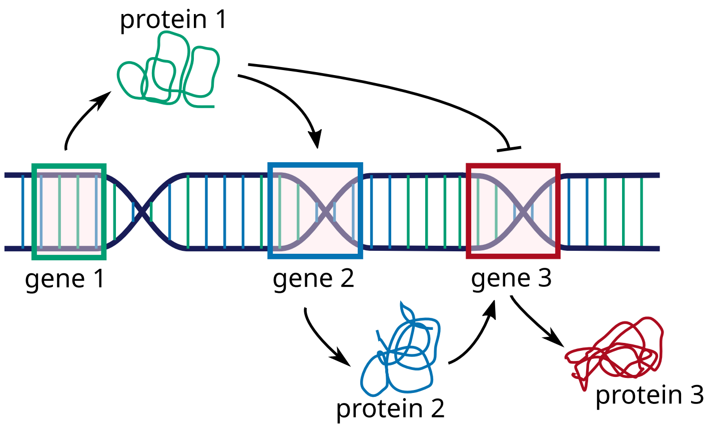
Loïc Paulevé

CNRS/LaBRI, Bordeaux, France

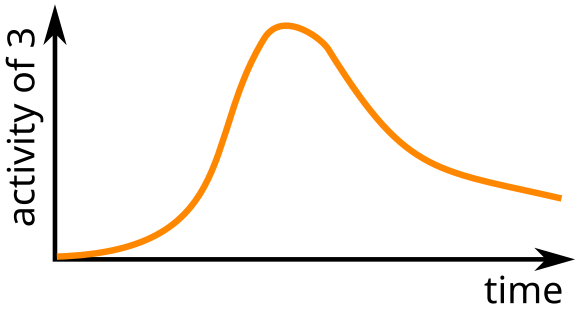
https://loicpauleve.name

Joint work with S. Haar, T. Chatain, J. Kolčák (Inria Saclay/LSV)

Boolean networks and systems biology

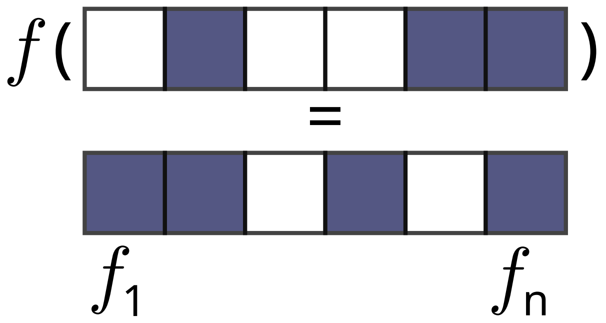


$$f_1(\mathbf{x}) = 1$$
$$f_2(\mathbf{x}) = \mathbf{x}_1$$
$$f_3(\mathbf{x}) = \text{not } \mathbf{x}_1 \text{ and } \mathbf{x}_2$$



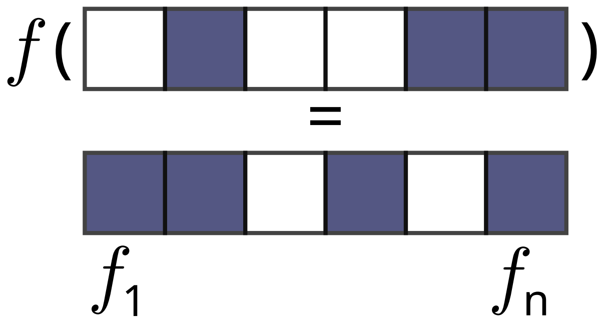
- BNs widely used as "mechanistic" models of biological processes (cellular differentiation, tumorigenesis, cycles, ...)
- Validation of BN models: reproduce observed behaviors
- Reachability ~ observed change of state of components
- Attractors ~ stable behaviors (possibly sustained oscillations)

Boolean Network (BN) $f : \mathbb{B}^n \to \mathbb{B}^n$

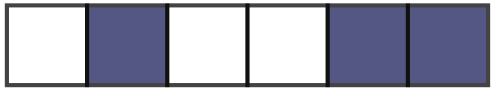


Discrete dynamical system with semantics specifying how a configuration of the network evolves in time

Boolean Network (BN) $f : \mathbb{B}^n \to \mathbb{B}^n$

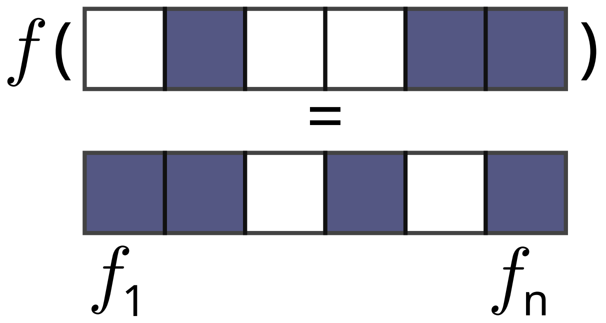


Configuration: $x \in \mathbb{B}^n$

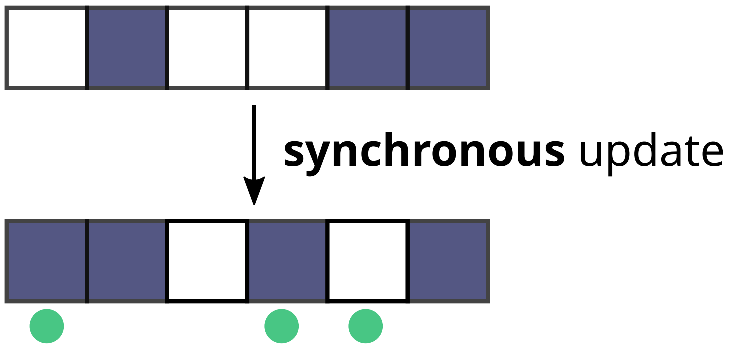


Discrete dynamical system with semantics specifying how a configuration of the network evolves in time

Boolean Network (BN) $f : \mathbb{B}^n \to \mathbb{B}^n$

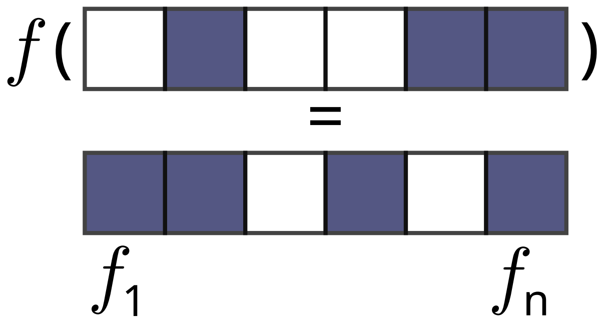


Configuration: $x \in \mathbb{B}^n$

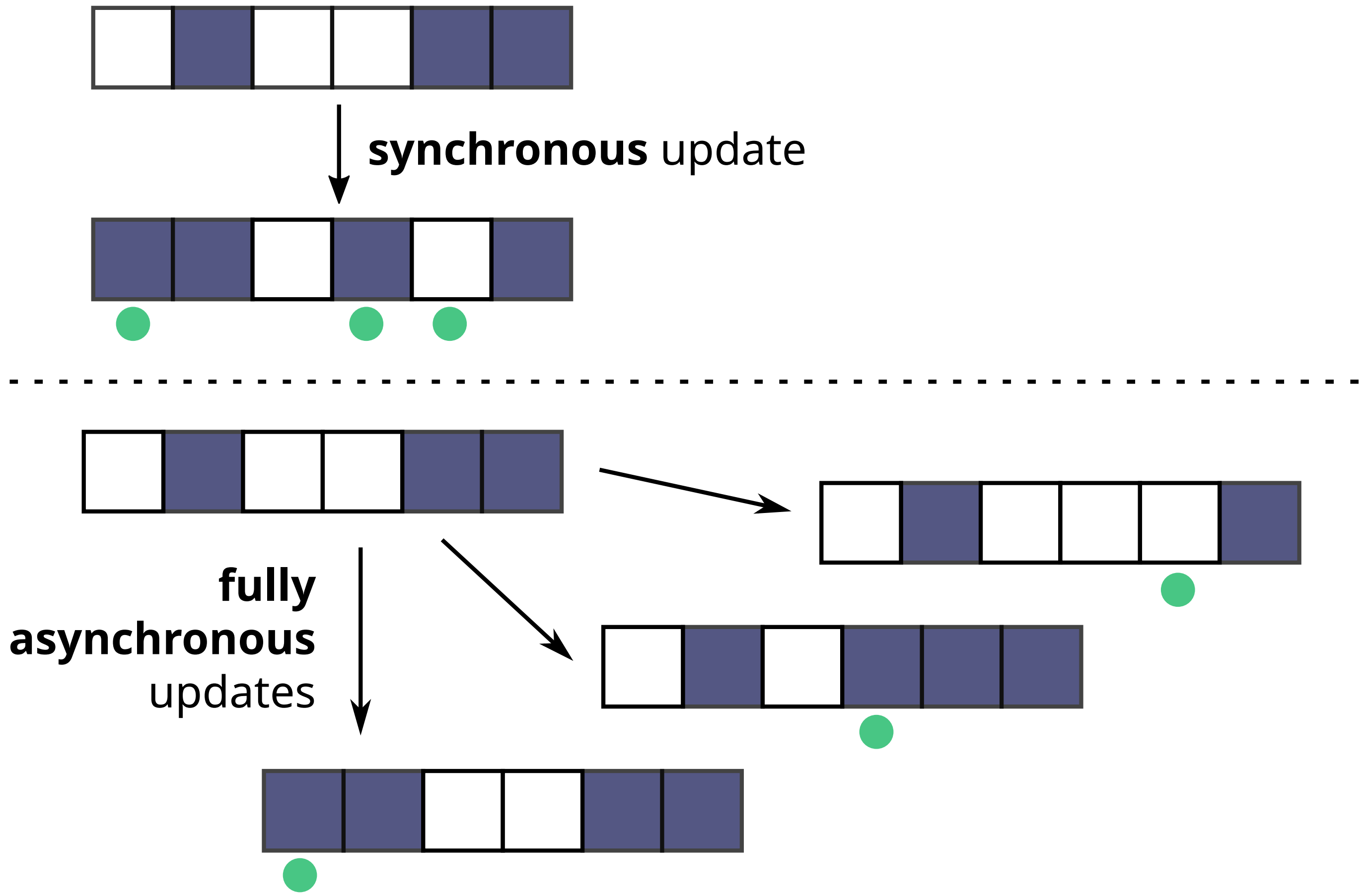


Discrete dynamical system with semantics specifying how a configuration of the network evolves in time

Boolean Network (BN) $f : \mathbb{B}^n \to \mathbb{B}^n$

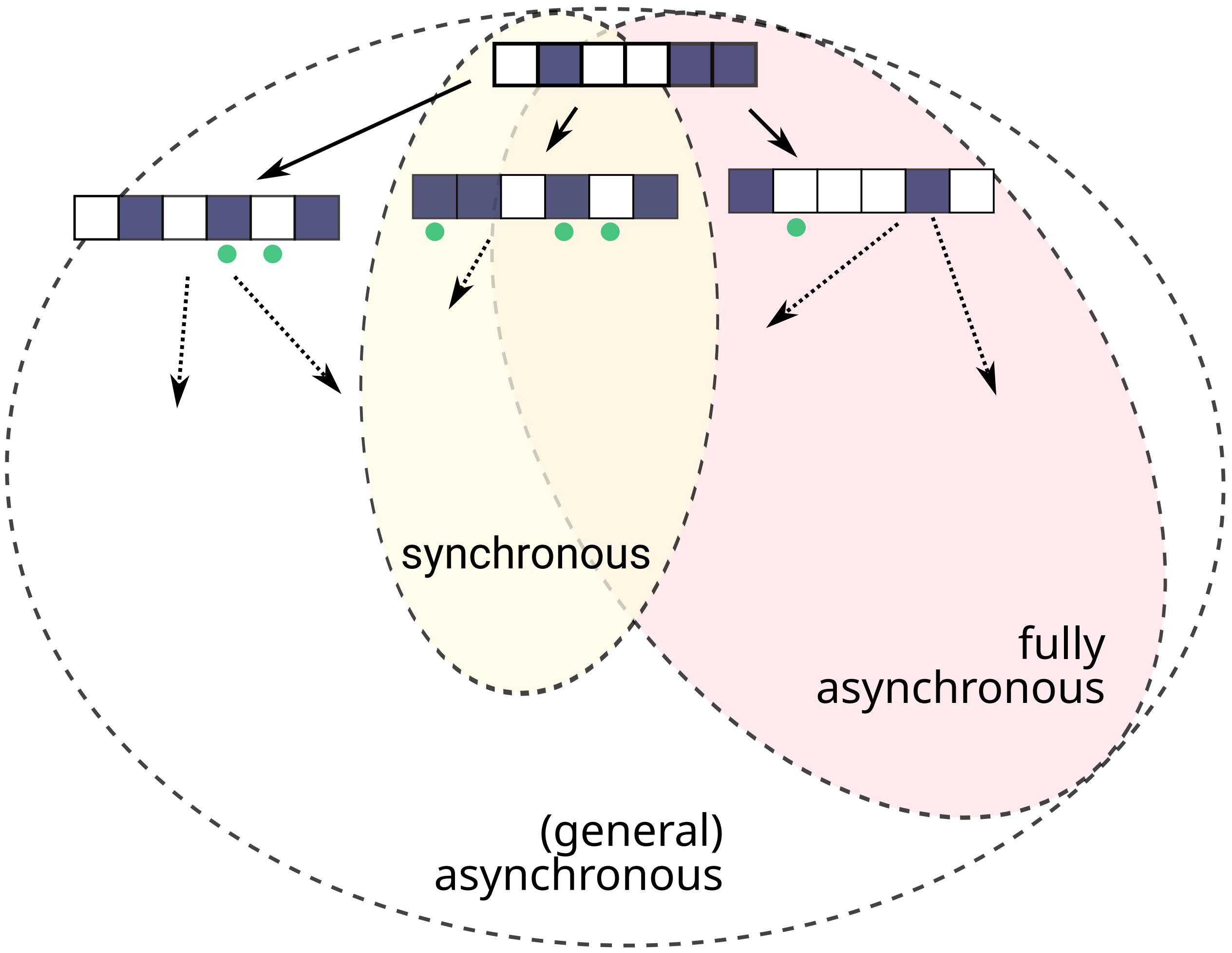


Configuration: $x \in \mathbb{B}^n$



Discrete dynamical system with semantics specifying how a configuration of the network evolves in time

Reachable configurations



- Given BN f and semantics σ

$$\rho^f_\sigma(x) \subseteq \mathbb{B}^n$$

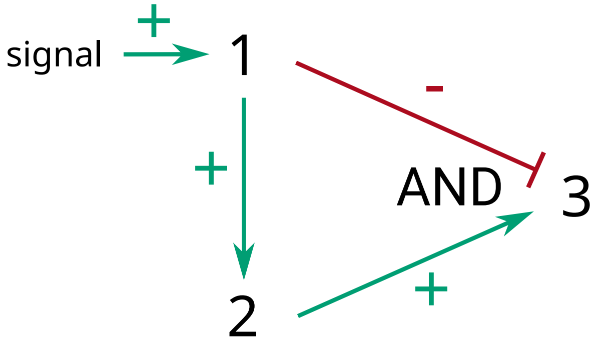
is the set of configurations reachable from configuration x

- For reachability, general asynchronous semantics also includes sequential, bloc-sequential, ...

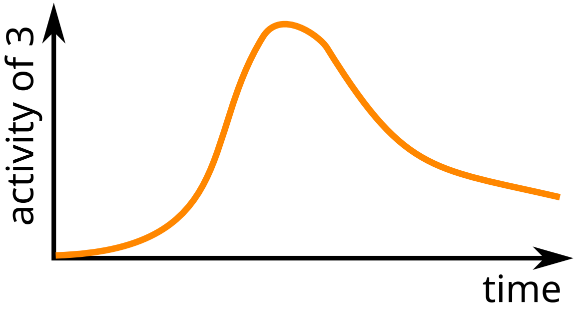
but is it complete?
(w.r.t. what?)

Boolean modeling of the I3-FFL (Incoherent feed-forward loop)

Regulation motif



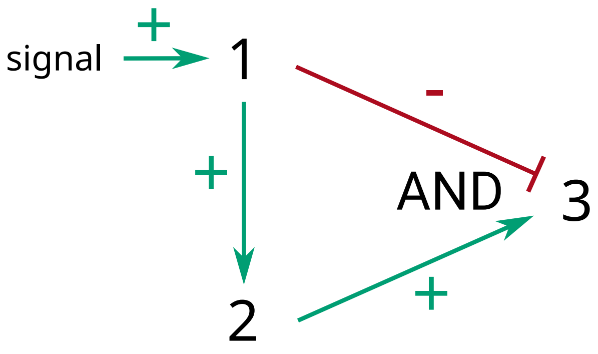
Observed output



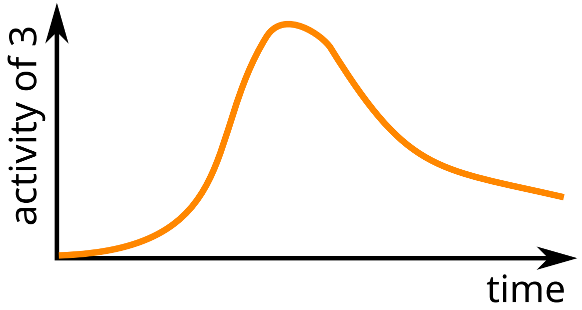
with quantitative models, and
synthetically designed DNA circuits

Boolean modeling of the I3-FFL (Incoherent feed-forward loop)

Regulation motif



Observed output



with quantitative models, and
synthetically designed DNA circuits

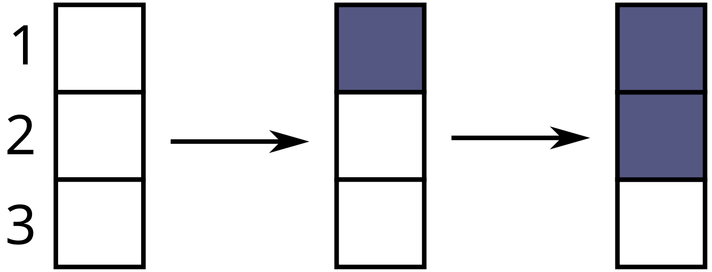
Boolean network

$$f_1(\mathbf{x}) = \text{signal}$$

$$f_2(\mathbf{x}) = \mathbf{x}_1$$

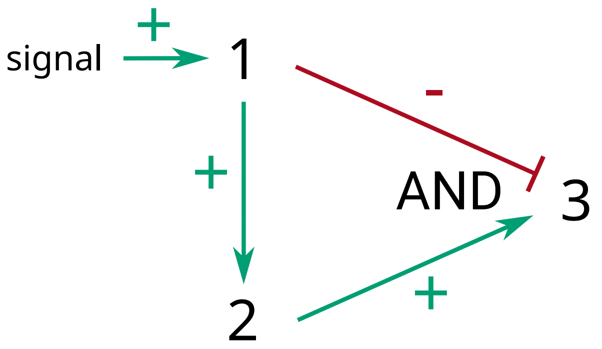
$$f_3(\mathbf{x}) = \text{not } \mathbf{x}_1 \text{ and } \mathbf{x}_2$$

Asynchronous dynamics from 000

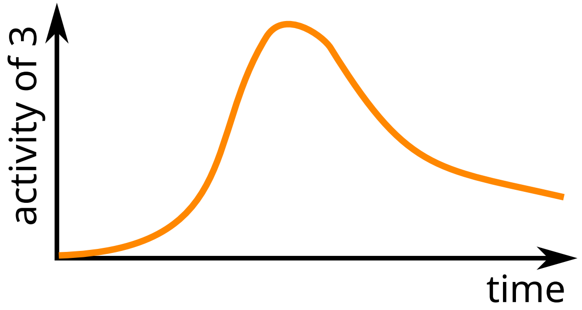


Boolean modeling of the I3-FFL (Incoherent feed-forward loop)

Regulation motif



Observed output



with quantitative models, and
synthetically designed DNA circuits

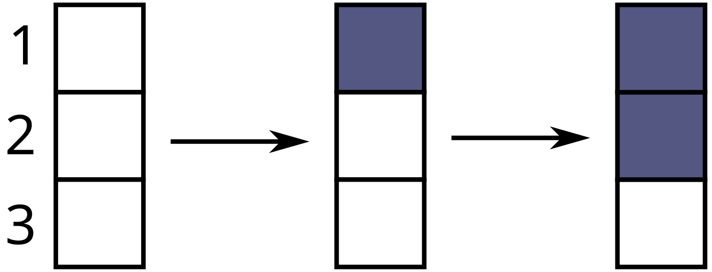
Boolean network

$$f_1(\mathbf{x}) = \text{signal}$$

$$f_2(\mathbf{x}) = \mathbf{x}_1$$

$$f_3(\mathbf{x}) = \text{not } \mathbf{x}_1 \text{ and } \mathbf{x}_2$$

Asynchronous dynamics from 000



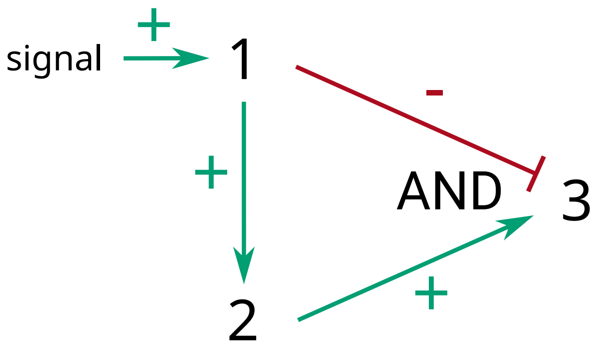
➥ **impossible to activate 3...**

- model validation fails but the logic is correct!
- no BN matching the motif works..

➥ incoherent abstraction for reachability...

Boolean modeling of the I3-FFL (Incoherent feed-forward loop)

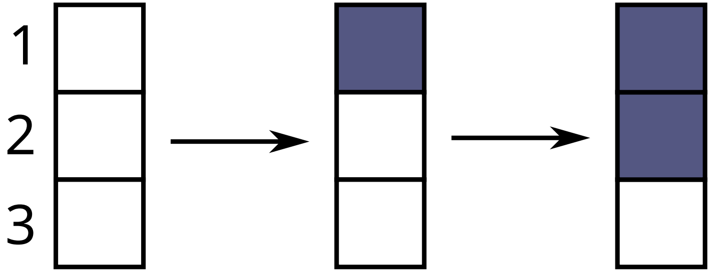
Regulation motif



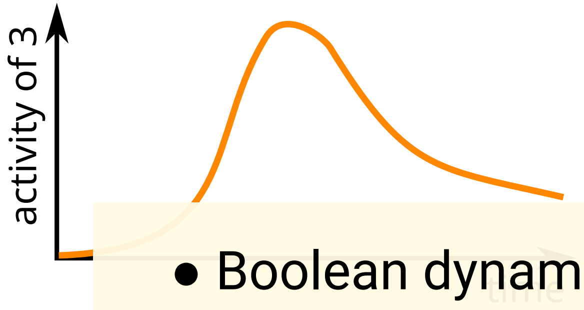
Boolean network

$$f_1(\mathbf{x}) = \text{signal}$$
$$f_2(\mathbf{x}) = \mathbf{x}_1$$
$$f_3(\mathbf{x}) = \text{not } \mathbf{x}_1 \text{ and } \mathbf{x}_2$$

Asynchronous dynamics from 000



Observed output



➥ impossible to activate 3...

- model validation fails but the logic is correct!

- no BN matching the motif works..

➥ incoherent abstraction for reachability...

with qua... synthe...

- Boolean dynamics fails to capture the period when 1 is high enough to activate 2, but not high enough to inhibit 3...
- one can fix the issue with multivalued networks, or delays

➥ adds many parameters, limiting their general application

(A)synchronous Boolean Networks

Bad abstractions of non-binary systems

- can miss behaviors...

... also includes stochastic methods

- impact reachable attractors: one can wrongly conclude an attractor is not reachable

Costly to analyze

- reachability and attractor properties are PSPACE-complete
- usually limited to 50-100 components; then requires approximations..

(A)synchronous Boolean Networks

Bad abstractions of non-binary systems

- can miss behaviors...
... also includes stochastic methods
- impact reachable attractors: one can wrongly conclude an attractor is not reachable

Costly to analyze

- reachability and attractor properties are PSPACE-complete
- usually limited to 50-100 components; then requires approximations..

Most Permissive Boolean Networks (MPBNs)

Paulevé et al, Nature Communications, 2020

Complete abstraction

- guarantees not to miss any behavior achievable by a quantitative model following the same logic
- remains stringent enough to capture differentiation processes

Highly scalable

- reachability: P/P^{NP};
attractor: $coNP/coNP^{coNP}$
- benchmarks with 100,000 components
- unlocks large-scale BN inference

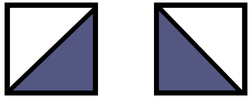
No additional parameters!

Most Permissive semantics - with pseudo dynamic states

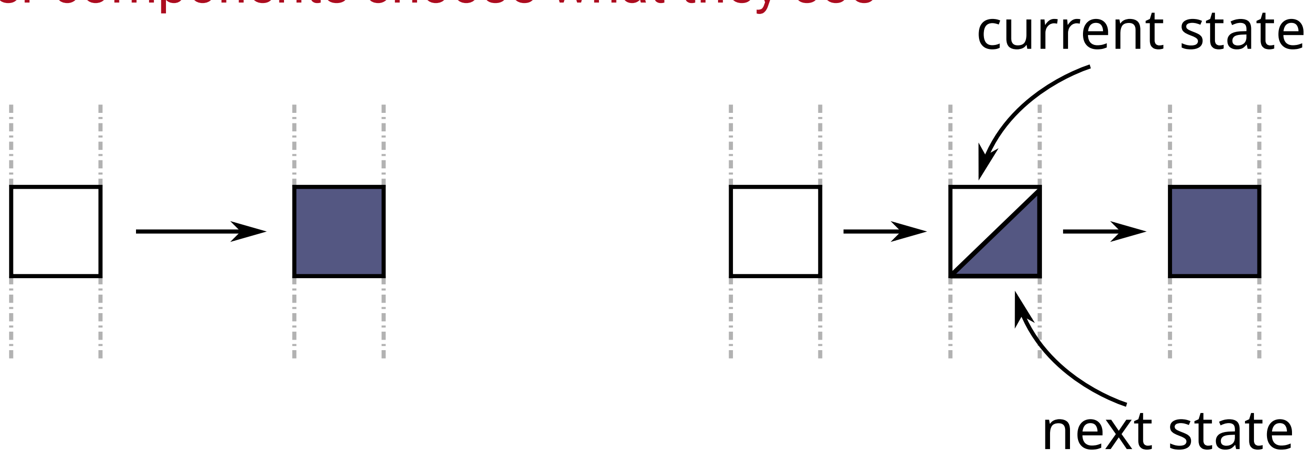
Two key ingredients:

- delay between firing and application of state change

  ↪ allow interleaving other state changes

- in pseudo "dynamic" states 

  other components choose what they see

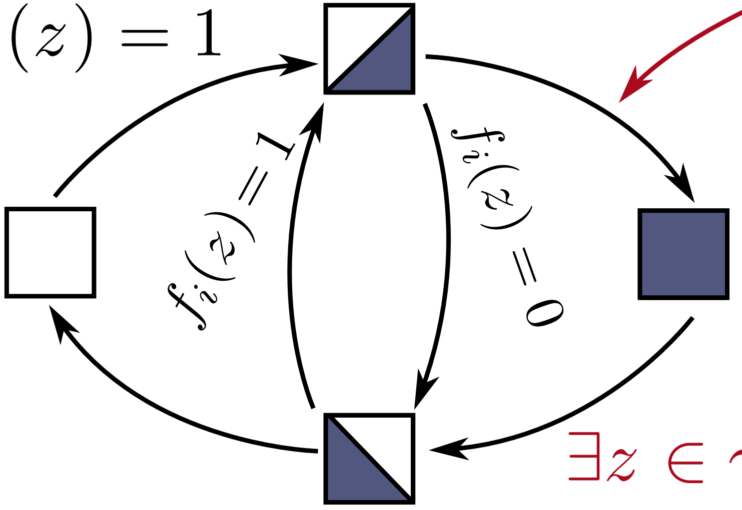


Most Permissive semantics - with pseudo dynamic states

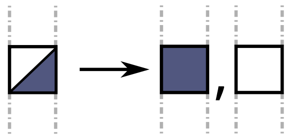
Automaton of component i

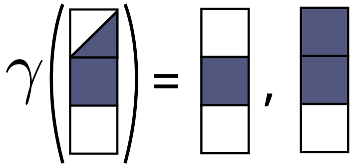
can fire anytime

$\exists z \in \gamma(x) : f_i(z) = 1$



$\exists z \in \gamma(x) : f_i(z) = 0$

$\gamma :$ 

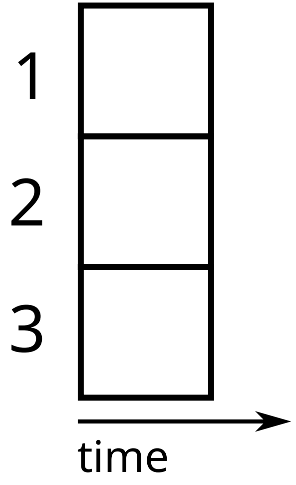
example: γ 

+ full-asynchronous interleaving

$$\rho^f_{\mathrm{mp}}(\mathbf{x}) := \{\mathbf{y} \in \mathbb{B}^n \mid \mathbf{x} \xrightarrow[\mathrm{mp}]{f}{}^{*}\, \mathbf{y}\}$$

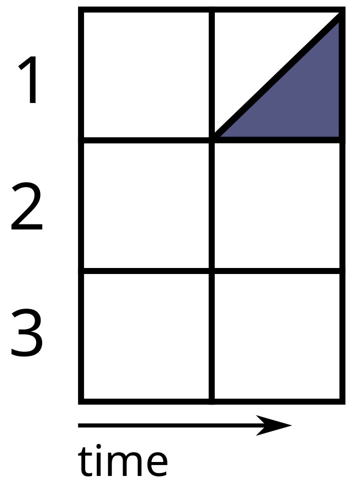
Most Permissive semantics - example of trajectory

$$f_1(\mathbf{x}) = \text{signal}$$
$$f_2(\mathbf{x}) = \mathbf{x}_1$$
$$f_3(\mathbf{x}) = \text{not } \mathbf{x}_1 \text{ and } \mathbf{x}_2$$



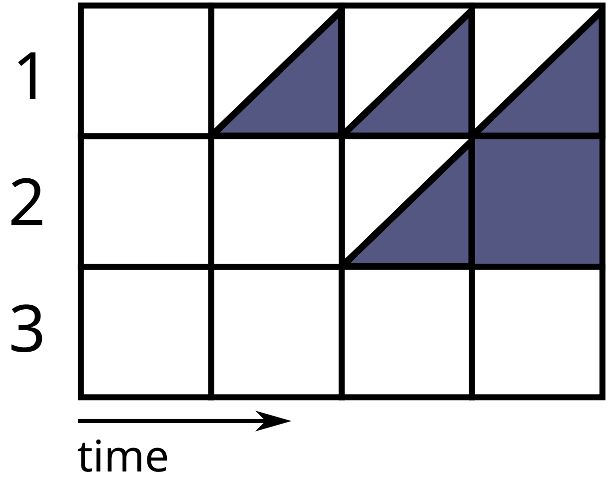
Most Permissive semantics - example of trajectory

$$f_1(\mathbf{x}) = \text{signal}$$
$$f_2(\mathbf{x}) = \mathbf{x}_1$$
$$f_3(\mathbf{x}) = \text{not } \mathbf{x}_1 \text{ and } \mathbf{x}_2$$



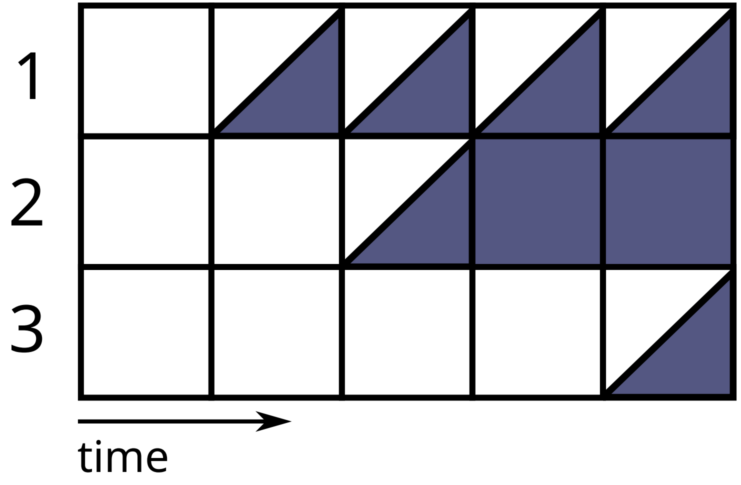
Most Permissive semantics - example of trajectory

$$f_1(\mathbf{x}) = \text{signal}$$
$$f_2(\mathbf{x}) = \mathbf{x}_1$$
$$f_3(\mathbf{x}) = \text{not } \mathbf{x}_1 \text{ and } \mathbf{x}_2$$



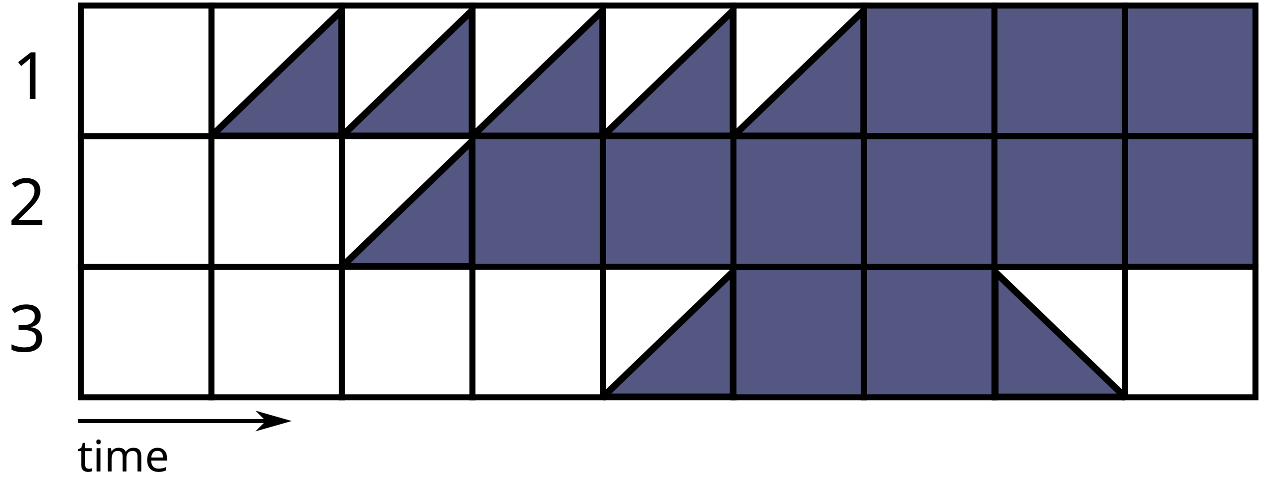
Most Permissive semantics - example of trajectory

$$f_1(\mathbf{x}) = \text{signal}$$
$$f_2(\mathbf{x}) = \mathbf{x}_1$$
$$f_3(\mathbf{x}) = \text{not } \mathbf{x}_1 \text{ and } \mathbf{x}_2$$



Most Permissive semantics - example of trajectory

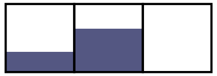
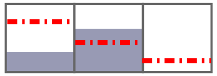
$$f_1(\mathbf{x}) = \text{signal}$$
$$f_2(\mathbf{x}) = \mathbf{x}_1$$
$$f_3(\mathbf{x}) = \text{not } \mathbf{x}_1 \text{ and } \mathbf{x}_2$$

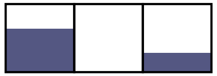
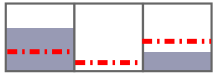
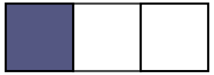


Most Permissive semantics - completeness and minimality

Completeness

A multivalued network $F : \mathbb{M}^n \to \{-1, 0, 1\}^n$
is a refinement of a Boolean network *f* iff

$F_i($  $) > 0 \implies \exists$  $: f_i($  $) = 1$

$F_i($  $) < 0 \implies \exists$  $: f_i($  $) = 0$

Most Permissive semantics simulates
any multivalued refinement with asynchronous update

↪ extends to ODEs

Minimality

- $\mathbf{y} \in \rho^f_{\mathrm{mp}}(\mathbf{x}) \implies \exists$ MN *F* refining *f* : $m \cdot \mathbf{y} \in \rho^F_{\mathrm{async}}(m \cdot \mathbf{x})$

Most Permissive semantics - complexity

Reachability problem

given configurations $x, y \in \mathbb{B}^n$
decide wether

$$y \in \rho^f_\sigma(x)$$

P with locally-monotonic BNs
P^{NP} in general
(there is always a MP trajectory of linear length between reachable configurations)

Attractor

Non-empty set of configurations $A \subseteq \mathbb{B}^n$
s.t. $\forall x \in A, \rho^f_\sigma(x) = A$

In-attractor problem

Given a configuration $x \in \mathbb{B}^n$
decide wether it belongs to an attractor

coNP with locally-monotonic BNs
$coNP^{coNP}$ in general
(MP attractors are minimal trap spaces)

$f : \mathbb{B}^n \to \mathbb{B}^n$ is *locally monotonic* whenever
$\forall i \in \{1, \dots, n\}, \exists \preceq^i \in \{\leq, \geq\}^n : \forall \mathbf{x}, \mathbf{y} \in \mathbb{B}^n,$
$(\mathbf{x}_1 \preceq^i_1 \mathbf{y}_1 \wedge \dots \wedge \mathbf{x}_n \preceq^i_n \mathbf{y}_n) \Rightarrow f_i(\mathbf{x}) \leq f_i(\mathbf{y})$

Most Permissive Boolean Networks - in practice

Attractors

- Fixed points are identical to (a)synchronous BNs
- Attractors are the minimal trap spaces of *f*;
- Less but bigger attractors than with (a)synchronous BNs

Applications to models of differentiation from litterature

- Recover the same predictions for reachable attractors
- Low computational cost: no need for approximations with model reduction
- Access to nature of attractors in large models

➥ Formally guaranteed to capture behaviors that only multivalued discrete models could capture with (a)synchronous interpretations
➥ If MPBNs cannot reach an observation, no quantitative refinement can do it

Software

mpbn Python library - https://**github.com/pauleve/mpbn**

- reachability and (reachable) attractors in locally-monotonic BNs
- shipped in the CoLoMoTo Docker http://colomoto.org/notebook
- based on Answer-Set Programming (clingo)

```
bn = mpbn.load("model.bnet")
bn.reachability(x, y)  # True if there is a trajectory from x to y
list(bn.attractors())   # List all attractors
list(bn.attractors(reachable_from=x))   # all attractors reachable from x
```

BoNesis - https://**github.com/bioasp/bonesis** (work in progress)

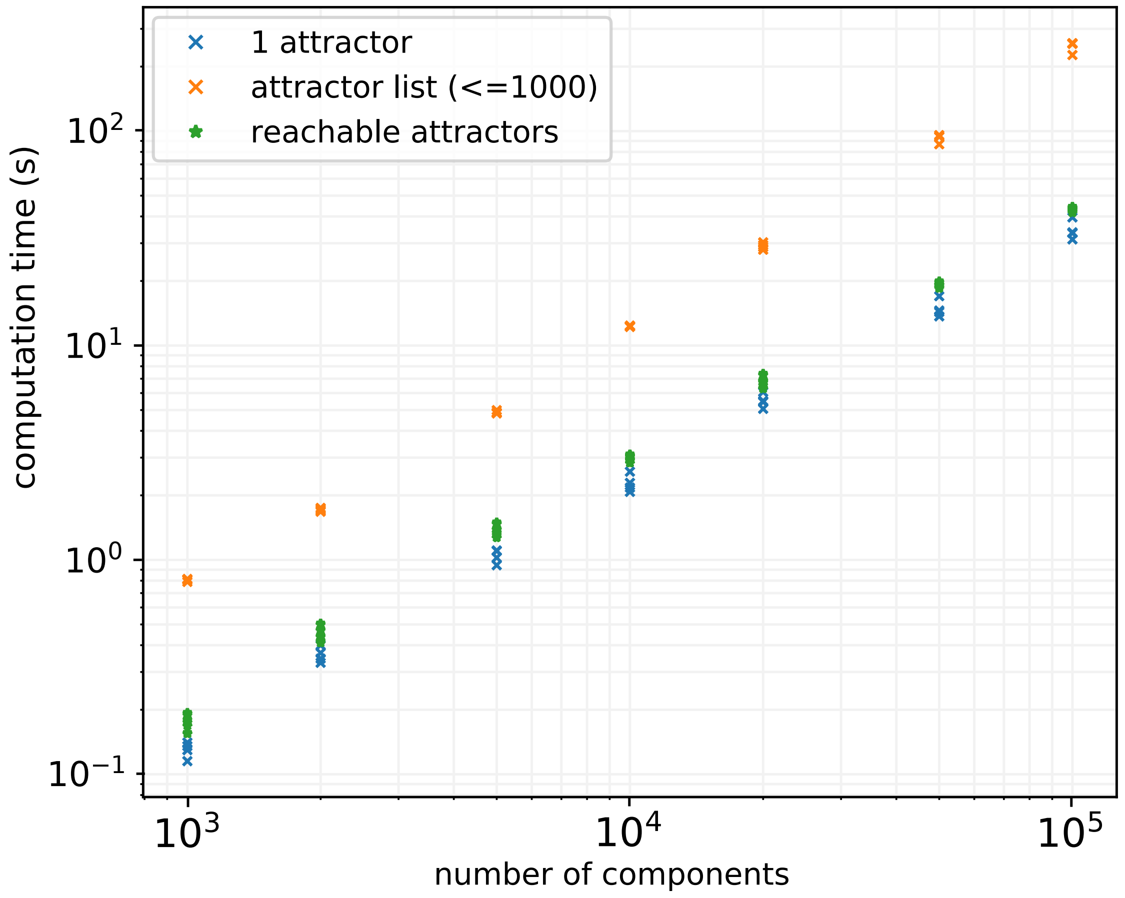
- MPBN synthesis (PhD work of S Chevalier)
- can be used to verify advanded properties on a given BN with MP semantics

Scalability experiments

Random scale-free inhibitor-dominant BNs; in-degree up to 1,400

Time for computing
- one attractor
- up to 1,000 attractors
- reachable attractors from a random initial configuration

"VLBNs" (Very Large Boolean Networks)
doi:10.5281/zenodo.3714876
Notebooks at:
doi:10.5281/zenodo.3936123



Application to models of cell fate decision

- MPBNs predict more trajectories: can they still predict losses of attractor rechability? (key feature of differentiation/fate decision models)

- Model of tumour invasion by Cohen et al 2015 (32 components)

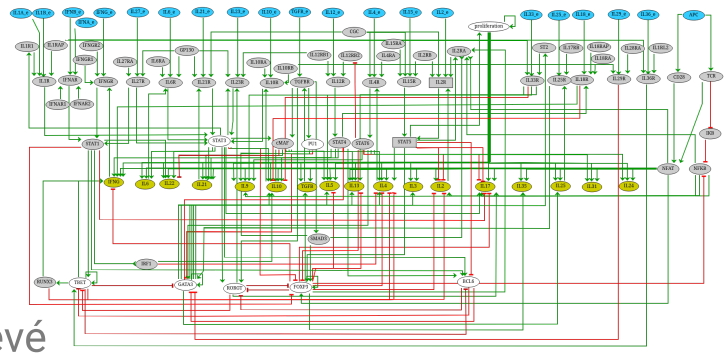
↪ same predictions on reachable attractors in different mutation settings

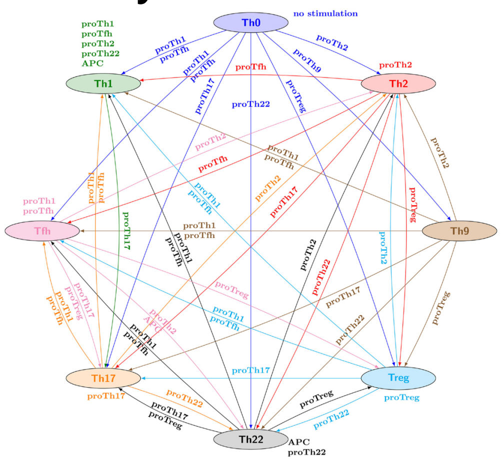
- Model of T-Cell differenciation by Abou-Jaoude 2015 (101 components)

↪ same predictions for the reprogramming graph

↪ applied directly on the large scale model (original study used reduced one)

↪ access to the nature of attractors





Conclusion

(A)synchronous semantics

- difficult to justify for mechanistic models (inconsistent abstractions)
- can miss important behaviors: lead to reject valid models of biological systems
- have limited tractability (state space explosion)

Most Permissive semantics

- correct abstraction: multilevel/quantitative refinemnts only remove behaviours
- simpler complexity: genome-scale tractability

Missing features (work in progress)

- Quantification of reachable attractors (like MaBoSS with full-async BNs)
- How to represent MP dynamics? transition graph is not adequat..